Reprinted from JOURNAL OF EXPERIMENTAL PSYCHOLOGY,
Vol. 45, No. 3, March, 1953
Printed in U. S. A.

STUDIES OF DISTRIBUTED PRACTICE: VIII. LEARNING AND RETENTION OF PAIRED NONSENSE SYLLABLES AS A FUNCTION OF INTRALIST SIMILARITY

BENTON J. UNDERWOOD[1]

Northwestern University[2]

The rate of learning paired-adjective lists is unrelated to intertrial intervals up to 2 min. when lists are presented at a 2:2-sec. rate, i.e., 2 sec. for the stimulus alone and 2 sec. for the stimulus and response appearing together (8). With materials of low meaningfulness the data are somewhat contradictory. In one study Hovland (4) found that a 2-min. rest after each trial resulted in no faster learning than did a 6-sec. rest for paired nonsense syllables. In a subsequent study (5) however, the same conditions did produce more rapid learning with spaced practice than with massed. Furthermore, the differences were magnified if pairs were presented at a 1:1-sec. rate. No resolution of these conflicting data is available.

Early studies by Hovland (e.g., 3) have consistently shown that in serial learning distributed practice facilitates acquisition. Previously in his writings, therefore, Hovland had theoretically tied the phenomenon of facilitation by distribution to remote associations which are formed in serial learning. But, having found that facilitation will occur with spaced practice of paired nonsense syllables, his last theoretical position removes the tie to serial associations and postulates what seems to be a work-inhibition theory. This theory (5) appears to be one which simply states that the more work done per unit of time the greater the likelihood that distribution will facilitate learning. Considerable evidence could be marshalled in support of this position. If one is to hold to such a theory, however, specification of the relationship between work and other variables should be suggested. One such variable is meaningfulness. That meaningfulness is in some way a variable in determining whether or not spaced practice facilitates is shown by the fact that learning of serial adjectives is not influenced as much by spaced practice as is learning of serial nonsense lists (9,12). Also, the learning of short verbal-discrimination lists of adjectives presented at a 1:1-sec. rate

[1] R. L. Morgan and E. J. Archer supervised the gathering of the data; Mr. Archer and Jack Richardson are largely responsible for the statistical analysis.

[2] This work was done under Contract N7onr-45008, Project NR 154-057, between Northwestern University and The Office of Naval Research.

is not influenced by spaced practice (13). Of course, to handle such findings, one could merely postulate that the greater the meaningfulness, the less the work required.

Another variable which may have to be related to work theory, or at least incorporated into some theory of distribution, is intratask similarity. In the first place, Hovland's discard of the need for remote associations in learning before distribution will facilitate learning is not easy to accept. One must first inquire into the basic processes which produce remote associations. One defensible position is that remote associations represent generalization tendencies. If this is so, and if generalization tendencies exist in paired-associate learning (as seems likely, e.g., 8), the quantitative continuity between serial and paired-associate learning can be maintained on the basis of the argument that no essentially different processes are represented by the two forms of learning. It is thus still plausible that the key to a theory of distributed practice lies in number of generalization tendencies (produced by intratask similarity) and not in the amount of work. Again, however, work theory could simply state that the greater the generalization the greater the work.

Accumulating evidence suggests that no simple theory of work inhibition will be entirely satisfactory for both learning and retention phenomena associated with spaced practice. It has been shown that while distribution of practice of paired-adjective lists does not affect acquisition rates, differences in retention occur as a function of the intertrial rests during learning (8). In another study (11) intralist similarity was varied in serial nonsense lists. The results showed that facilitation occurred with distributed practice but this facilitation was no greater with lists of high similarity than with those of low similarity. If one did wish to postulate that work and intralist similarity are directly related, such evidence is difficult to handle. And in this same experiment, although intralist similarity produced wide differences in rate of learning, errors did not increase systematically with intralist similarity as one would expect by generalization theory. In brief, we are short on acceptable theory to account for the few well-established facts we have and we are short on empirical laws of some generality describing relationships between intertrial rest, related variables, and learning-retention phenomena.

The first purpose of the present experiments is to provide additional facts on the effect of intertrial rest in learning paired-associate nonsense lists. As pointed out above, Hovland's data are contradictory on this matter. The second purpose is to get information on the relationship between intralist similarity and intertrial interval as they affect acquisition. The third purpose is to measure retention as a function of the two variables.

Procedure

General.—There were five separate experiments. Within each experiment (in which similarity is constant) there are three conditions of distribution, namely, 4, 30, and 60 sec. between each trial. Similarity differs among experiments, with intralist stimulus similarity and intralist response similarity varying independently. In Exp. I both stimulus similarity and response similarity are low. In *Exp. IIs* intralist stimulus similarity is medium and response similarity low. In Exp. IIIs, intralist stimulus similarity is high and response similarity is low. In Exp. IIr, intralist response similarity is medium and stimulus similarity low. In Exp. IIIr, intralist response similarity is high, and stimulus similarity low. Actually, the same lists were used in Exp. IIr and IIIr as in IIs and IIIs, simply by "turning the lists over" so that what were formerly stimuli now became responses, and the

former responses now became stimuli. To summarize:

Experiment	*Stimulus Similarity*	*Response Similarity*
I	Low	Low
IIs	Medium	Low
IIIs	High	Low
IIr	Low	Medium
IIIr	Low	High

Lists.—All lists were constructed of nonsense syllables of 46.67% to 53.33% association value according to Glaze (2). Each list consisted of ten pairs. In the above conditions low similarity means that there was no repetition of any consonant among stimulus (or response) items and no vowel was used more than three times. In medium similarity each of five letters was used twice as the first letter of a syllable and each of five different letters was used twice to end a syllable. In high similarity only three letters were used to start the ten syllables (two letters started three and one started four). Likewise, only three (although different) consonants were used to end syllables. In no case was a vowel (center letter) used more than three times in a list and commonly each occurred twice. Thus, for a low stimulus (or response) series, 20 consonants were used to make up the 10 items; for medium similarity 10 were used, and for high similarity, 6. To avoid stimulus-response generalization similarity between stimuli and responses was kept as low as possible. Obviously some repetition was unavoidable but the basic principle followed was that no letter used to start a syllable on the stimulus side was used to start a syllable on the response side. Repetition did occur between the first letters of stimulus syllables and last letters of response syllables or between first letters of response syllables and last letters of stimulus syllables. Also, repetition between lists was kept as low as possible since each experiment required three lists, one for each distribution interval. Finally, one practice list (of medium similarity) was used for all five experiments. All lists were presented on a Hull-type drum at a 2:2-sec. rate with learning by the anticipation method, in which *S* spelled the syllable. Three orders of presentation were used to avoid serial learning.

Specific conditions.—Each experiment employed 36 college students as *Ss*. Each *S* had one practice session and four experimental sessions, the last experimental session requiring only recall and relearning of the list learned the previous day. On the practice day *S* learned the list to 6 out of 10 correct responses on a single trial, the learning being by massed practice (4 sec. between trials). He was then instructed concerning the rest-interval activity to be used to fill the intertrial rests during distributed practice. This activity, symbol cancellation, has been explained in detail in a previous report (10). Following these instructions *S* continued learning with a 30-sec. rest between each trial until one perfect trial was achieved. After 5 min., recall and relearning took place. On experimental days *S* learned three lists, one under each of the three intertrial rest conditions (4, 30, 60 sec.), but, of course, only one list each day. Learning was always carried to one perfect recitation and recall and relearning (by massed practice) occurred after 24 hr. A given list was always recalled and relearned before learning the list which in turn was recalled and relearned the following day.

With three conditions and three lists, complete counterbalancing of both lists and conditions is achieved with 36 *Ss*. The groups of *Ss* were matched on the common practice list.

Results

Practice list.—The mean number of trials to learn the practice list was 29.75, 29.06, 29.42, 28.72, and 30.33 for Exp. I, IIs, IIIs, IIr, and IIIr, respectively. The variance is homogeneous for the distributions on which these means are based and *F* is less than 1. For all groups combined the product-moment correlation between trials to learn the practice list and trials to learn the three experimental lists combined is .57 ± .07. The mean number of errors per trial on the practice list was 1.23, 1.41, 1.07, 1.24, and 1.33 for the five groups in order. The *F* is 1.56 and with 4 and 175 *df* an *F* of 2.42 is needed for significance at the 5% level of confidence. The correlation between error frequency on practice list and on three experimental lists combined was .65 ± .07. It may be concluded that the groups were initially comparable.

Learning of experimental lists.—The mean number of trials to learn the experimental lists for the three intertrial rest intervals for all five experiments is shown in Table 1 and plotted in Fig. 1. The statistical analysis initially deals with Exp. I, IIs, and

TABLE 1

MEAN TRIALS TO LEARN PAIRED NONSENSE LISTS AS A FUNCTION OF STIMULUS AND RESPONSE SIMILARITY AND LENGTH OF INTERTRIAL REST*

Exp.	Intertrial Rest		
	4 sec.	30 sec.	60 sec.
I	23.97	22.42	23.33
IIs	24.78	23.44	28.28
IIIs	29.86	32.89	29.39
IIr	24.19	21.42	21.50
IIIr	25.22	23.69	24.27

* Estimate of σ_m is 1.30.

IIIs independently, and then with Exp. I, IIr, and IIIr. In this and all other analyses to be presented, variance was shown not to be heterogeneous.

The essential terms of analysis of variance for Exp. I, IIs, and IIIs are shown in Table 2. The rationale and proof of this technique have been presented elsewhere (1). For evaluation of significance of similarity the proper error term is Ss/Similarity. The F of 5.29 is beyond the value (4.82) needed for significance at the 1% level. For evaluating the effects of other variables the appropriate error term is Pooled Ss × Practice/Similarity. Terms necessary for evaluating list differences within a given level of similarity are not included. It will be observed that intertrial rest is not a significant source of variance, F being less than 1. In Fig. 1 the curves for Exp. IIs and IIIs (medium and high stimulus similarity) have quite different shapes. This is reflected in the interaction term in Table 2 (Intertrial Rest × Similarity) which is significant at about the 3% level of confidence. The rather bizarre shapes for these two curves argue against accepting this interaction as being psychologically meaningful. From the analysis, however, two conclusions are straightforward for these three experiments in which interlist stimulus similarity was manipulated: (*a*) as interlist stimulus similarity increases, rate of learning decreases, and (*b*) intertrial rest produces no differences in learning for any level of similarity.

Turning next to results produced by variation in response similarity (Exp. I, IIr, and IIIr), we may first note that similarity is not an effective variable as far as trials to learn is concerned. The complete statistical analysis will not be presented. It is sufficient to report that F for similarity is less than 1. Thus, variation in similarity which produces differences in rate of acquisition when it obtains among stimuli has little influence on learning when present among responses. Although Fig. 1 might suggest that learning was more rapid under conditions with 30- and 60-sec. intertrial rests than for 4-sec. rest for Exp. I, IIr, and IIIr, the F for intertrial rest (2.10) falls considerably short of the value (3.04) needed for the 5%

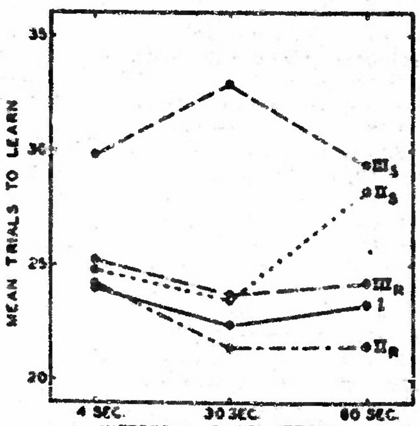

FIG. 1. Learning as a function of intertrial interval and intralist similarity. I, IIs, and IIIs indicate increasing stimulus similarity; I, IIr, and IIIr indicate increasing response similarity.

level. Again, then, we have no evidence to support Hovland's (5) finding that distributed practice will facilitate acquisition of paired nonsense lists.

Finally, it must be pointed out that the lists in which response similarity was manipulated were learned more rapidly than those in which stimulus similarity was manipulated. Disregarding intertrial rest, the mean number of trials to learn the medium stimulus-similarity lists (Exp. IIs) was 25.50. When these lists were "turned over" so that similarity was among responses the mean was 22.37 (Exp. IIr). The corresponding values for the high-similarity lists (Exp. IIIs and IIIr) were 20.71 and 24.40. *F* is highly significant (5.36). The analysis of errors during learning, to which we now turn, adds information which will be useful in interpreting these differences.

Errors during learning.—The mean number of overt errors per trial in learning is shown for each condition in Fig. 2. The first fact demonstrated by Fig. 2 is that variation in response similarity (IIr and IIIr) resulted in greater error frequency than did variation in stimulus similarity (IIs and IIIs). Thus, while learning was more rapid with variation in response similarity than with variation in stimulus similarity, more errors per trial were made in the former instance.

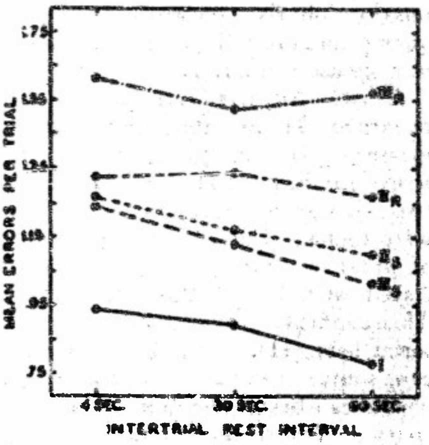

FIG. 2. Overt errors per trial in learning as a function of intertrial interval and intralist similarity

A second fact is that similarity was a significant source of variance for Exp. I, IIr, and IIIr (*F* is 13.2) with error frequency varying directly with similarity. With variation in stimulus similarity *F* is 4.15, which falls between the 1% and 5% levels. Here there is no direct relation between similarity and error frequency since the number of errors for the high-similarity lists (IIIs) is slightly less than for the lists of medium similarity (IIs).

A third fact to be noted in Fig. 2 is that for Exp. I, IIs, and IIIs, error frequency decreases with increasing length of intertrial rest. The *F* is 8.31 which is well beyond the 1% confidence level. On the other hand, no

TABLE 2

ANALYSIS OF VARIANCE FOR MEAN NUMBER OF TRIALS TO LEARN IN EXP. I, IIs AND IIIs

Source of Variation	*df*	Mean Square	*F**
Similarity	2	1586.0402	5.29
Ss/Similarity	105	299.8429	
Practice	2	884.0216	14.84
Practice × Similarity	4	3.0679	—
Intertrial rest	2	21.5772	—
Intertrial rest × Similarity	4	177.3040	2.98
Lists/Similarity (pooled)	6	——**	
Pooled Ss × Practice/Similarity	198	59.5680	
Total	323		

* With 2 and 108 (105) *df*, *F* at .05 level is 3.09, and at .01, 4.82. With 2 and 200 (198) *df*, *F* at .05 level is 3.04, and at .01, 4.71. With 4 and 200 (198) *df*, *F* at .05 is 2.41, and at .01, 3.41.

** This mean square has no meaning except when broken down to show differences in list difficulty within each similarity level.

relation obtains between error frequency and intertrial rest for variation in response similarity.

Recall.—All lists were recalled and relearned 24 hr. following original mastery. In analyzing the recall scores for Exp. I, IIs, and IIIs, only two significant sources of variance were found. One of these was stage of practice, with poorer recall associated with later stages of practice. This confirms previous findings with serial lists (11). The second significant source of variance was similarity, with recall better the higher the similarity. Intertrial interval during learning had no influence on retention as measured by recall.

When Exp. I, IIr, and IIIr were subjected to analysis of variance, only stage of practice was found to be significant. Unlike variation in stimulus similarity, variation in response similarity produced no differences in recall. It will be remembered that variation in stimulus similarity produced differences in rate of learning whereas response variation did not. It is possible, therefore, that differences in recall as a function of stimulus similarity may be due to differences in number of reinforcements (correct responses) during original learning. To examine this possibility an item analysis of learning has been made in which number of reinforcements is held constant. When this is done, differences in recall as a function of stimulus similarity disappear. The results of this analysis for Exp. I, IIs, and IIIs are shown in Fig. 3. When items having the same number of reinforcements are compared (ignoring intertrial interval), there are no appreciable differences in recall as a function of similarity. The raw recall scores also showed better retention for Exp. IIs and IIIs than for IIr and IIIr. Here again, however, when compared by item analysis so that frequency of reinforcement is equal, no differences of any consequence are apparent. It must be concluded that differences in similarity per se have little influence on retention as measured by recall. This confirms previous results with paired adjectives (8).

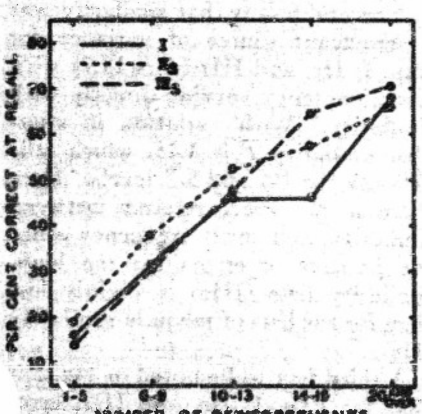

FIG. 3. Recall as a function of stimulus similarity. The base line indicates the number of correct responses made for an item during original learning.

It was pointed out earlier that more errors occurred in learning lists in which response similarity varied than in those in which stimulus similarity varied. It might be expected that such differences would occur in overt errors at recall and such is the case. For Exp. IIs and IIIs the total errors for all three conditions was 150 and 104 respectively; for Exp. IIr and IIIr the corresponding values were 171 and 186.

Relearning.—Taken singly, none of the major variables influences rate of relearning. However, for both stimulus similarity and response similarity the interaction between stage of practice and similarity is highly significant. The F for stimulus similarity is 6.02 and for response similarity, 5.33, with an F of 4.71 needed for the 1% level.

With high similarity (either IIIs or IIIr), the mean number of trials to relearn decreases as a function of stage of practice, the values being 9.33, 8.28, and 6.06 for Exp. IIIs and 8.02, 7.36, and 6.89 for IIIr, for the three stages of practice in order. For low similarity (Exp. I), on the other hand, the mean number of trials to relearn increases as a function of stage of practice, these means being 5.33, 6.08, and 7.58. This interaction between similarity and stage of practice may be expected on the basis of differences in *interlist* similarity which resulted from variation in *intralist* similarity. In these experiments three experimental lists were used for each level of similarity. Intralist similarity was varied by manipulating the number of repeated consonants. With low intralist similarity consonants became repeated among lists as a result of avoiding duplication within lists. With high similarity repetition among consonants in different lists was avoided because of high repetition within lists. Therefore, we would expect considerable interlist interference for low similarity lists and amount of interference should be directly related to stage of practice, i.e., number of previous lists learned. Thus, it would seem that with low-similarity lists, interference among lists more than counteracted the effects of practice, while this was not true in the high-similarity lists. Other measures (e.g., recall scores) are consistent with this hypothesis but since this finding is secondary to the major purpose of the paper, such data will not be detailed here.

Finally, it may be mentioned that the errors per trial during relearning followed a pattern almost identical with that during original learning.

Discussion

In none of the five experiments has any facilitation resulted in learning by spaced practice as compared with massed. Furthermore, the results argue against a simple work-inhibition type theory. In Hovland's (5) study in which facilitation by distributed practice did occur, the mean number of trials to learn was 16.4 for the massed condition and 13.5 for the spaced condition. In the present experiments the mean number of trials varied from 22 to 33. If work inhibition is involved, it would be suspected that greater inhibition would have occurred with massed practice in our experiments than in Hovland's experiments. Now, it is true that Hovland's distributed condition consisted of a 2-min. intertrial rest whereas our longest interval was 1 min. But, some consistent trend toward faster learning with distributed practice should be present in our data even with 1-min. rests if work inhibition develops. Such was not the case. Another possibility to account for the discrepancy may be contained in the fact that Hovland's *Ss* were extremely well practiced. This is reflected in the mean values given above. In a previous study (9) we have shown that for serial learning differences between rate of learning under massed and distributed practice do not change appreciably as a function of stage of practice but it is still remotely possible that stage of practice could be a variable for paired-associate learning. For the present the contradiction between Hovland's two studies (4,5) and our results cannot be satisfactorily resolved.

Confirming previous findings (3), the present results give no support to generalization as a critical process in learning paired-associate lists by massed and distributed practice. With variation in stimulus similarity wide differences in

rate of learning occurred but no facilitation by spaced practice was observed for any level of similarity. There was, then, no interaction between similarity and intertrial rest. Similarly, in learning serial nonsense lists (11) interaction between these two variables has not been significant. However, interaction between intertrial rest and similarity did appear in an earlier study (12) using serial adjectives, with lists of high similarity showing more facilitation than lists of low similarity. This finding is being checked at the present time with a more extensive series of conditions. Except for this one case, variation in intratask similarity does not seem to reveal processes which are essential to the study of the influence of distributed practice.

The data have shown that while variation in stimulus similarity among paired associates produces significantly different rates of learning, comparable variation among response terms does not. That similarity among responses has an influence is shown by the direct increase in overt errors with increase in similarity. It seems likely that with increasing similarity among responses not only are interference tendencies increased but also factors making for facilitation in learning. It has been shown elsewhere (7) that in a transfer situation with variation in response similarity between lists, large amounts of positive transfer may occur even with a high frequency of errors from the first to the second list. In the present experiments with increase in response similarity the number of different letters to be given by *S* in spelling the syllables decreases. Thus, as similarity increases, *S* must remember fewer individual letters (which should facilitate acquisition) but must discriminate which particular combinations of these few letters must be attached to the different stimuli (which should retard learning). If these two processes are in rough balance, little change in rate of learning as a function of response similarity should be expected. On the other hand, with increase in stimulus similarity only the inhibiting or interfering factor increases. *S* has to learn to respond with dissimilar responses to increasingly similar stimuli. As similarity among these stimuli increases, discrimination among the stimuli to which the discrete responses are to be attached becomes more difficult. In this situation, however, it is not clear why errors do not increase directly with increase in stimulus similarity. The fact that the same result was evident in serial learning of nonsense syllables (11) shows that the law has some generality but no process has been suggested to account for it.

The final result requiring comment is the decrease in errors as a function of intertrial rest; with variation in stimulus similarity overt errors decreased directly with length of intertrial rest. With variation in response similarity no such relationship was observed. It may also be reported that in a comparable series of experiments (manuscript in preparation) with paired adjectives, about the same results were obtained except that some decrease in errors (significant at 5% level) was present with variation in response similarity. Such findings might be interpreted as evidence for a differential-forgetting theory as proposed by McGeoch (6). This theory states that erroneous response tendencies, since they are weaker than correct response tendencies, will drop out or be forgotten more rapidly over a rest interval. If this is the case in the present experiments, we would expect faster learning by distributed practice (unless some other process counteracts). The theory has been offered to account for faster learning by distributed practice than by massed; here we find evidence for the theory in terms of errors, but there is no facilitation by distributed practice. Error frequencies for the present experiments have been analyzed by Vincent-type curves as a function of stage of learning a given list. The massed and distributed conditions were plotted separately. This analysis showed no appreciable differences in the shapes of the curves as a function of intertrial interval although there was a trend for the differences in error frequency to be greater early in learning than late in learning. If it is reasonable to suppose that weak error tendencies are more frequent early in learning than late in learning, then these error curves could be said to support a differential-forgetting theory. On the other hand, the fact that the error frequencies during relearning by massed practice were less for lists learned by distribution than for lists learned by massing would argue against the differential-forgetting theory unless some permanent error-reduction process is postulated. It must be repeated again, however, that no differences in learning, recall, or relearning were found so that if the differential-forgetting theory is said to be supported by the error data, it must be quickly added that no noticeable effect on correct responses was observed at any time. In previous experiments with serial lists (10) and in the present experiments, error frequency seems to bear little relationship to performance measured by correct responses.

Another factor which might be relevant to error frequencies is the rest-interval activity. It was demonstrated earlier (10) that color naming during the distribution interval produces many more errors than does symbol cancellation (the activity used in the present experiments). Furthermore, the same differences were present during relearning by massed practice. These differences were interpreted to be the conse-

quence largely of a responding set induced by color naming although we were unable to rule out entirely an error-depressant effect of symbol cancellation. In symbol cancellation very few errors are made; instructions for accuracy are given each day *S* cancels. It is possible that *S* develops a tendency to make fewer errors in learning as a consequence of actually making very few errors in symbol cancellation. The present data are not well suited for an internal analysis by stage of practice to determine if such a set is developing. Such an analysis has been made and the results show trends which would support the set hypothesis but the differences do not attain statistical significance. Moreover, if such an error-reducing set is built up, it is not a general finding since it has not been present in serial learning (11) nor has it been evident in the present experiments in which response similarity was manipulated. For the present, the correlates of the phenomenon of fewer errors with longer intertrial rests must be left unspecified.

Summary

Five experiments were performed to study the effect of (*a*) intertrial rest and (*b*) intralist stimulus and intralist response similarity on learning and retention of lists of paired nonsense syllables. The three intertrial intervals in each experiment were 4, 30, and 60 sec. In one experiment the lists had low stimulus and low response similarity; in two others the stimulus similarity was medium and high, respectively, with response similarity low; in the other two experiments the response similarity was medium and high, respectively, with stimulus similarity low. Similarity was manipulated by varying the frequency of repeated letters making up the syllables of a given list. All learning was carried to one perfect trial with retention of each list measured after 24 hr. Each experiment involved 36 *S*s. The results show:

1. There was no influence of intertrial rest on rate of learning in any of the five experiments. Overt error frequency was inversely related to intertrial rest for experiments in which stimulus similarity was manipulated.

2. Difficulty of learning increased with increase in stimulus similarity but remained unchanged for variation in response similarity. Overt errors increased directly with response similarity but not with stimulus similarity.

3. No differences in retention occurred as a function of either intertrial rest or similarity.

The results do not support Hovland's latest study in which learning of paired-nonsense lists was facilitated by distributed practice. The major differences in procedure are that Hovland used a 2-min. intertrial rest (compared to the maximum 1-min. rest in the present experiments) and his *S*s were much better practiced than those used here. The present findings further support the conclusion that intratask similarity is not an important variable in the study of the influence of intertrial rests on learning. The fact that variation in stimulus similarity produced differences in rate of learning whereas response similarity did not may be explained on the basis of the different roles played by similarity in the two cases. The finding that fewer errors occurred with distributed practice than with massed practice when stimulus similarity is varied is unaccounted for, but such differences in errors appear to have no relationship to rate of learning.

(Received August 18, 1952)

References

1. Archer, E. J. Some greco-latin analysis of variance designs for learning studies. *Psychol. Bull.*, 1952, 49, 321–337.
2. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, 35, 255–269.
3. Hovland, C. I. Experimental studies in rote-learning theory: III. Distribution of practice with varying speeds of syllable presentation. *J. exp. Psychol.*, 1938, 23, 172–190.

4. HOVLAND, C. I. Experimental studies in rote-learning theory: V. Comparison of distribution of practice in serial and paired-associate learning. *J. exp. Psychol.*, 1939, 25, 622–633.
5. HOVLAND, C. I. Experimental studies in rote-learning theory: VIII. Distributed practice of paired associates with varying rates of presentation. *J. exp. Psychol.*, 1949, 39, 714–718.
6. MCGEOCH, J. A. *The psychology of human learning*. New York: Longmans, Green, 1942.
7. UNDERWOOD, B. J. Associative transfer in verbal learning as a function of response similarity and degree of first-list learning. *J. exp. Psychol.*, 1951, 42, 44–53.
8. UNDERWOOD, B. J. Studies of distributed practice: II. Learning and retention of paired-adjective lists with two levels of intra-list similarity. *J. exp. Psychol.*, 1951, 42, 153–161.
9. UNDERWOOD, B. J. Studies of distributed practice: III. The influence of stage of practice in serial learning. *J. exp. Psychol.*, 1951, 42, 291–295.
10. UNDERWOOD, B. J. Studies of distributed practice: VI. The influence of rest-interval activity in serial learning. *J. exp. Psychol.*, 1952, 43, 329–340.
11. UNDERWOOD, B. J. Studies of distributed practice: VII. Learning and retention of serial nonsense lists as a function of intra-list similarity. *J. exp. Psychol.*, 1952, 44, 80–87.
12. UNDERWOOD, B. J., & GOAD, D. Studies of distributed practice: I. The influence of intra-list similarity in serial learning. *J. exp. Psychol.* 1951, 42, 125–134.
13. UNDERWOOD, B. J., & VITERNA, R. O. Studies of distributed practice: IV. The effect of similarity and rate of presentation in verbal-discrimination learning. *J. exp. Psychol.*, 1951 42, 296–299.